

2

NAVAL POSTGRADUATE SCHOOL

Monterey, California

AD-A251 936



DTIC
ELECTE
JUN 17 1992
S C D

THESIS

A COMPARATIVE EVALUATION
OF COMPUTER ACCESS CONTROLS

by

Timothy B. Pence

March, 1992

Thesis Advisor:

Moshe Zviran

Approved for public release; distribution is unlimited

92-15528

92 6 15 067

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE				
1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b RESTRICTIVE MARKINGS	
2a SECURITY CLASSIFICATION AUTHORITY			3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.	
2b DECLASSIFICATION/DOWNGRADING SCHEDULE				
4 PERFORMING ORGANIZATION REPORT NUMBER(S)			5 MONITORING ORGANIZATION REPORT NUMBER(S)	
6a NAME OF PERFORMING ORGANIZATION Naval Postgraduate School		6b OFFICE SYMBOL (If applicable) 37		7a NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c ADDRESS (City, State, and ZIP Code) Monterey, CA 93943-5000			7b ADDRESS (City, State, and ZIP Code) Monterey, CA 93943-5000	
8a NAME OF FUNDING/SPONSORING ORGANIZATION		8b OFFICE SYMBOL (If applicable)		9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c ADDRESS (City, State, and ZIP Code)			10 SOURCE OF FUNDING NUMBERS	
			Program Element No	Project No
			Task No	Work Unit Accession Number
11 TITLE (Include Security Classification) A COMPARATIVE EVALUATION OF COMPUTER ACCESS CONTROLS				
12 PERSONAL AUTHOR(S) Pence, Timothy B.				
13a TYPE OF REPORT Master's Thesis		13b TIME COVERED From To		14 DATE OF REPORT (year, month, day) 1992, March
				15 PAGE COUNT 77
16 SUPPLEMENTARY NOTATION The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.				
17 COSATI CODES			18 SUBJECT TERMS (continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUBGROUP		
			Computer Security; User Identification; Passwords	
19 ABSTRACT (continue on reverse if necessary and identify by block number) This thesis reports the results of a study which tested participants' abilities to recall five different types of computer passwords. Each participant was assigned in a randomized procedure to one of six response intervals. Recall testing of computer-generated passwords, user-created passwords, passphrases, associative passwords and cognitive passwords was conducted using a computer program which simulated system log-on procedures. This study indicates the relative merits of these five password types are more difficult to distinguish when data are collected in the realistic setting of a log-on simulation instead of via paper surveys, as was done in previous research				
20 DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS REPORT <input type="checkbox"/> DTIC USERS			21 ABSTRACT SECURITY CLASSIFICATION	
22a NAME OF RESPONSIBLE INDIVIDUAL Prof. Moshe Zviran			22b TELEPHONE (Include Area code) (408) 646-3094	
			22c OFFICE SYMBOL AS/Zv	

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted
All other editions are obsoleteSECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Approved for public release; distribution is unlimited

A Comparative Evaluation of Computer Access Controls

by

Timothy B. Pence
Lieutenant, United States Navy
B.S., United States Naval Academy, 1984

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN INFORMATION SYSTEMS

from the

NAVAL POSTGRADUATE SCHOOL
(March, 1992)

Author:

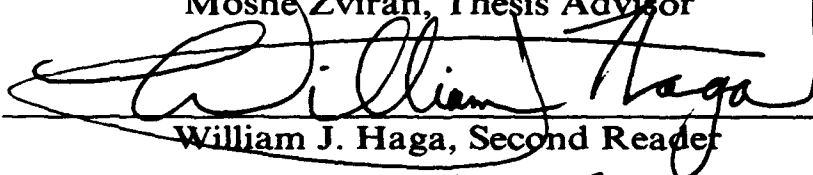


Timothy B. Pence

Approved by:



Moshe Zviran, Thesis Advisor



William J. Haga, Second Reader



David R. Whipple, Chairman
Department of Administrative Sciences

ABSTRACT

This thesis reports the results of a study which tested participants' abilities to recall five different types of computer passwords. Each participant was assigned in a randomized procedure to one of six response intervals. Recall testing of computer-generated passwords, user-created passwords, passphrases, associative passwords and cognitive passwords was conducted using a computer program which simulated system log-on procedures. This study indicates the relative merits of these five password types are more difficult to distinguish when data are collected in the realistic setting of a log-on simulation instead of via paper surveys, as was done in previous research.



Accession For	
NTIS GRANT	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

TABLE OF CONTENTS

I. INTRODUCTION	1
A. THE NEED FOR COMPUTER SECURITY	1
B. USER AUTHENTICATION MECHANISMS	2
1. Biometric Devices	3
2. Security Token Methods	7
3. Passwords	9
II. USER IDENTITY VERIFICATION WITH PASSWORDS	11
A. INTRODUCTION	11
B. THE IMPORTANCE OF SECURE PASSWORDS	11
1. Selecting Secure Passwords	12
C. TYPES OF PASSWORDS TESTED	13
1. Computer-generated Passwords	13
2. User-created Passwords	14
3. Passphrases	15
4. Associative Passwords	16
5. Cognitive Passwords	17
D. SUMMARY	18
III. RESEARCH METHODOLOGY	19
A. BACKGROUND	19
B. METHODOLOGY	20
1. Data Collection Description	20
2. Testing Recall of Passwords	28
IV. FINDINGS	35
A. COMPARISON OF SIMULATED LOG-ON SUCCESSES	35
1. Recall of Associative and Cognitive Passwords	36
2. Recall success versus log-on attempts	37
3. Recall of cognitive passwords	38
B. METHODS OF RECALL	40
1. Computer-generated passwords	40
2. User-created passwords	41
3. Passphrases	42
C. EASE OF PASSWORD RECALL	42
D. EASE OF PASSWORD USE	43
E. COMPUTER EXPERIENCE OF PARTICIPANTS	45
1. Length of prior computer experience	45
2. Types of computers used	46
V. DISCUSSION	47

A.	ANALYSIS OF SUCCESSFUL LOG-ON RATES	47
1.	Description of Analysis	47
2.	Results of Analysis	48
3.	Simulated Log-on Versus Individual Password Recall	49
4.	Benefits of Permitting Multiple Log-on Attempts	50
B.	ANALYSIS OF RECALL MECHANISMS	51
1.	Significance of Cognitive Password Recall ...	51
2.	How Secure Are Our Passwords?	52
C.	PERCEPTIONS VERSUS RECALL RESULTS ..	53
VI.	CONCLUSIONS	56
A.	THE "BEST" PASSWORD TYPE	56
B.	PAPER SURVEYS VERSUS COMPUTER STUDIES	56
1.	Comparison of Survey Methods	57
2.	Comparison with Previous Results	57
3.	Theory Versus Practice	58
	APPENDIX A	60
	APPENDIX B	62
	APPENDIX C	63
	LIST OF REFERENCES	68
	INITIAL DISTRIBUTION LIST	70

I. INTRODUCTION

A. THE NEED FOR COMPUTER SECURITY

In the immediate aftermath of the Persian Gulf War, U.S. Department of Defense (DoD) investigators found that computer hackers from the Netherlands were able to copy and modify data related to wartime U.S. military operations, as well as information on the transport of military equipment and personnel. Investigators reported the hackers gained access by using default passwords and exploiting flaws in computer operating systems (Alexander, 1991).

A study by the U.S. General Accounting Office found that 30% of the computer systems on Internet, a wide area computer network with thousands of subscribers, could be accessed by a password derived from a user identification, log-on, or identification spelled backwards (Salamone, 1991). Passwords based on log-on, user identification, or user name are vulnerable to "intelligent guessing" by would-be intruders. A password which appears in the dictionary (that is, a password which is an actual word) may be recovered through the use of a program which employs a computerized dictionary to rapidly guess tens of thousands of potential passwords.

The use of computers by governments, businesses and individuals continues to grow. Thousands of corporations, educational institutions and public agencies electronically link their mainframe

computer systems to promote efficiency through ease of information exchange. Additionally, a booming segment of the data processing market in recent years has been small, portable computers which can be used away from the office. A traveler who uses a laptop or notebook for on-the-go computing often uses the same computer to communicate via modem with the home office for end-of-the-day data dumps, electronic mail, or transmission of memos. Hence, accessibility of office computer systems has increased concurrent with the rapidly-growing portable computer market. While networking and interconnectivity of computer systems have numerous advantages, they provide an easy avenue for intruders to gain access to computer resources.

B. USER AUTHENTICATION MECHANISMS

A person who uses a computer without authorization may do so for a number of reasons. Among the simplest of these is theft of computing services. Many computer systems charge customers a fee based on usage; an uncertified user receives services for free. Sometimes the motivation for invasion of a computer system is malicious or mischievous. The intruder's intent may be to do damage in the former case or simply to experience the thrill of outsmarting the computer's security systems in the latter. Often the intruder seeks access to a computer's data for purposes of gaining information or modifying the data (Pfleeeger, 1989, pp. 11-13).

To prevent the loss, modification or compromise of data which can result when unauthorized persons are able to log-on to a computer, several user authentication mechanisms are available to system administrators.

1. Biometric Devices

Biometric authenticators use a person's physical traits to verify his/her identity. The many security tools in this category work in a similar manner: a biometric portrait of the subject is scanned or read by sensor devices, converted into digital data and stored. When an authorized user desires access to a protected computer, the trait used for authentication is tested and compared with the stored data (Alexander, 1990).

a. Handprint and Fingerprint Readers

Both handprint and fingerprint readers depend on the uniqueness of each individual's hand geometry or fingerprint ridges to identify him/her. Handprint readers also called palm readers measure the relative lengths of fingers when the hand is placed upon a template. Some models may scan lines on the palm of the hand. Because of the simple yet effective principle behind this design, handprint readers were the first type of biometric device to be made available on the commercial market (Parks, 1990).

Fingerprint readers scan to a finer degree than handprint readers and record measurements of the loops, whorls and arches that make up a single fingerprint. These devices are the lowest-cost option

for biometric security. A fingerprint reader can be purchased for as little as \$1000 (Alexander, 1990).

b. Voice Analyzers

A more elaborate biometric scheme involves testing a user's identity through voice recognition. A digitized pattern of an authorized user's voice is maintained by the computer. A typical scenario calls for the user to identify him/herself via the computer keyboard. The user then recites one or more words or phrases, which the computer compares with stored data, in order to gain access to the system. Such an authentication mechanism can be used when the user is physically located near the computer or can be used via phone lines (Penzias, 1990).

c. Retina Scanners

The pattern of blood vessels on the inside of an eyeball is unique for each person even identical twins. Retina scanners use this fact to verify a person's identity. A beam of low-intensity infrared light enters the eye through the pupil and scans a circular pattern upon the retina. A portion of the light is reflected back to a photodetector which records data at hundreds of points as the light beam traverses its arc. These data, a series of digitally-coded light intensity levels, are compared with future scans to authenticate a user requesting access (Fitzgerald, 1989).

d. Keystroke Analyzers

Among the more interesting concepts used to authenticate users is that of *keystroke latencies*, the elapsed time between keystrokes the user makes while using a computer keyboard. Research has shown that for repeatedly sampled strings of characters a person's keystroke pattern can be just as unique as a signature. The same muscles and neurological factors that form a signature are used for typing; it is therefore logical that each person types in a unique way that can be measured (Joyce and Gupta, 1990).

Employing this method, a new user to the computer system might be asked to repeatedly type his/her name or, for better security, a phrase of his/her own choosing for the authenticator software. A mean digital signature is then calculated from the several samples. The signature consists of the average latency between each successive keystroke. Future log-on attempts are then compared with the latency signature to validate the user (Joyce and Gupta, 1990).

e. Signature Analyzers

A person's signature has long been a customary means of identification for official matters. Methods exist which allow a computer to identify a person by examining the characteristics of his/her signature. One approach is to optically scan a signature written on an ordinary piece of paper; the scan results can be digitized and compared with future signatures. Unfortunately, a digital record

of the static image of a signature leaves the computer open to spoofing by skilled forgers (Mital and Lau, 1989).

A better means of recording a signature is through an examination of a person's handwriting dynamics. The pressure exerted on a piece of paper by the writing instrument as it is moved through the signature process is as unique as the signature itself. Furthermore, pressure variances are not visible during or after the signing process. This eliminates the forgery problem noted above.

In this method, an individual signs his/her name using a stylus on a pressure-sensitive pad. Varying pressure on the pad generates a voltage which is measured digitally. The pressure on the pad is sampled numerous times during the signing; the resulting plot of voltage versus time produces a pressure waveform characteristic of the individual's signature. This waveform can be compared against subsequent signatures in future log-on attempts (Mital and Lau, 1989).

A second method of signature dynamics measurement involves quantifying the writing instrument's motion as opposed to the pressure it exerts on the writing surface. A person signs his/her name with a pen which is wired to a port in the computer. During the signing, the pen's motion is tracked by piezoelectric accelerometers wired to it. In this way, the exact movement of the pen is recorded and can be compared against future signatures. (Fitzgerald, 1989).

f. Drawbacks

While biometric authentication eliminates the possibility of unauthorized log-on through compromise of a password, the methods discussed above have limitations. Changes in a person's physical characteristics or health can affect a test's outcome. A user who cuts his finger may not pass a fingerprint reader's scrutiny while another who sees her manicurist may alter the dimensions of her hand as seen by a handprint reader. A person who catches a cold may find a voice analyzer unable to recognize her speech while an amateur athlete suffering from tennis elbow might type or write differently and be unrecognized by a keystroke or signature analyzer. Finally, the purchase, installation and operation of these systems can be expensive for small businesses.

2. Security Token Methods

Rather than identifying a person by his/her physical characteristics, security tokens depend upon the possession of a device to verify a user is who he/she claims to be. Tokens can be employed by themselves to identify a computer user or they can be used to provide a third level of computer security in addition to the commonly-used log-on name and password. Security tokens have become more popular in recent years because of the growing number of people who use computers remotely via wide area network or modem (Wood, 1991).

a. Magnetic Cards

Probably the simplest security token technique calls for users to be issued a card which contains identifying information, usually on a magnetic stripe. The card is examined by a reading device; if it contains a signature the device recognizes, the bearer of the card is allowed access to the computer system. This application is most often used to control access to the *area* of computer terminals as opposed to individual terminals. To regulate each microcomputer or mainframe terminal individually requires each terminal have a card reader.

b. Smart Tokens

Other techniques escape the need for a magnetic reader. One method employs a device the size of a credit card which generates and displays a new password at some regular interval. An electronic clock in the card's microprocessor is synchronized with a similar clock in the host computer. When a user calls the host, he/she inputs a personal identification number and the card-generated password. Since both are required for a successful log-on, only an authorized user in possession of the smart token can gain access to the computer. Periodic regeneration of passwords prevents an intruder from making use of old passwords (Fitzgerald, 1989).

A similar procedure involves the use of a small calculator-like device. During log-on, the host computer displays a challenge number to the terminal which the user keys into the device. Using an

algorithm known to the host, the device calculates and displays a response. The user then inputs this in answer to the host's challenge number. After receiving the correct passcode, the host computer asks for a conventional user identification and password. The passcode thus provides a third level of security in addition to the other two log-on parameters (Wood, 1991).

c. Screen Readers

Another variation of the smart token procedure calls for the host computer to display on the user's screen a bar code like the ones used in supermarkets. The bar code challenge is scanned by a matchbook-sized token carried by an authorized user; the token displays a response number which the user inputs at the keyboard. A drawback to this method becomes apparent with the use of laptop and notebook computers. Some of these machines' displays lack the brightness to allow the bar code reader to accurately scan the code on the screen (Wood, 1991).

3. Passwords

Despite the availability of the computer security measures mentioned above, most computer systems which require user authentication still use a combination of user identification a user's name or assigned ID code and a password known to the user and the host computer. Passwords are the simplest way to incorporate security into a computer system. Software to enable their use is readily available or can easily be written from scratch. Password-

based authentication procedures are easy to use, and the cost of their administration is low. Care must be taken, however, in the creation and use of passwords to ensure they enhance system security.

II. USER IDENTITY VERIFICATION WITH PASSWORDS

A. INTRODUCTION

Because of their simplicity of use, low cost and ease of implementation passwords are the most widely employed means of user authentication in computer systems. Typically, a person who desires to log-on to a computer enters a portion of his/her name or an assigned user identification code along with a password. If the computer's log-on software verifies the identification and password match correctly with stored data, then the user is granted access. Depending upon the security requirements of the system, the user may be asked to provide additional passwords to access specific files, directories, procedures or application programs. Often, however, a user is given "carte blanche" access to a system's resources after correctly entering only one password.

B. THE IMPORTANCE OF SECURE PASSWORDS

Because a single password is frequently a computer system's only line of defense against intruders, as much effort as possible should go into selecting a password which will resist attempts at intrusion. A secure password should be impossible to guess *and* easy for the user to remember (Smith, 1991). Unfortunately, these two qualities are mutually exclusive to a degree. A random string of uppercase and

lowercase letters, numbers, and other keyboard symbols is very difficult to guess, but it is also difficult to remember. Users of such passwords often write them down so they won't be forgotten. This degrades the secrecy of the password and thus the computer security it provides.

1. Selecting Secure Passwords

Computer intruders often gain initial access to a system through intelligent guessing. A spouse's name, a portion of one's social security number, anniversary dates, birthdays, one's address all are examples of publicly-available information from which many computer users create passwords. Robert Morris, a designer of the Internet worm which caused damage to dozens of computer systems, has compiled a list of 73 words that can access at least one user on 90% of the large computer systems on Internet (Salamone, 1991). Computer users must be made aware of techniques which can dramatically improve the security of passwords they create.

Adherence to a few simple rules allows users to design customized, easily-remembered passwords that are also secure (Padovano, 1991).

- Include digits in the password
- Mix uppercase and lowercase letters
- Don't use a proper name or variation of a proper name
- Don't use a word found in a dictionary

- Don't use *QWERTY* keyboard patterns such as "asdfgh" or "a;sldkfj"

Two methods of password creation which typically result in hard-to-guess passwords involve combining two words or using the first letters of a multi-word phrase (Smith, 1991). For instance, a cooking enthusiast who also likes to vacation at Lake Tahoe might combine the words "chef" and "Tahoe" to create "chefTahoe". In an example of the second method, the same cooking enthusiast might create a password from the phrase "barbecued spare ribs with honey glaze sauce": "bbqsrwhgs". To further increase security, these methods can be applied in order to guarantee the inclusion of uppercase and lowercase letters as well as digits. For instance, the phrase "My friend Harriet has two children" might create the password "MfHh2-".

C. TYPES OF PASSWORDS TESTED

The study that is the basis for this thesis compared rates of recall of five different types of password mechanisms. Each is described below.

1. Computer-generated Passwords

Perhaps the simplest way to ensure a user employs a secure password is to arbitrarily assign one. A person who has no input in the creation of a password will not have the opportunity to create a personalized password vulnerable to intelligent guessing. Hence, some computer systems simply assign passwords to their users. Those

systems which use this method often utilize a computer program which produces passwords created from random alphanumeric characters. While this method yields passwords which are very secure, it has a large disadvantage in that users invariably find such passwords hard to remember. They therefore often resort to writing down the password as a memory aid; unfortunately, the act of writing down the password also degrades security.

Previous research (Beedenbender, 1990) found that although 13% of the people given a random, computer-generated password were able to remember it after a period of three months, 86% of them were able to do so only because they had written the password down. Better results were achieved when the computer-generated passwords were designed to be non-sensical but pronounceable non-dictionary words. In this case, the successful recall rate after three months was 38%. Of those who correctly remembered the password, 67% said they recalled it because it was pronounceable. Another 17% wrote it down.

2. User-created Passwords

Computer systems which allow users to construct their own passwords most frequently employ the user-created password as a means of identity verification. This is usually done with restrictions on password length (a minimum and maximum number of characters are specified) and on content (spaces and some non-alphanumeric keyboard characters may not be allowed). Additionally, the system's

password algorithm may not distinguish between uppercase and lowercase forms of the same letter.

User-created passwords are susceptible to the variety of vulnerabilities discussed in previous sections of this chapter. Previous research (Sawyer, 1990) has shown that users frequently choose passwords which are less than secure and seldom change them. In a survey of mainframe computer users who were allowed to create their own passwords, 65% reported their passwords were based on a meaningful detail in their lives such as a name or date. 80% used only alphabetic characters in their passwords. Despite being allowed to construct their own password, 20% of users admitted they still found it necessary to write the password down. Finally, 80% of those polled said they never changed their password, while an additional 15% said they changed passwords less frequently than once a year. It can thus be seen that, despite their popularity, user-created passwords can have a host of security weaknesses unless the means of their creation is carefully monitored.

3. Passphrases

A passphrase attempts to make a password harder to guess through sheer arithmetic. Passphrases have the same characteristics as passwords, but they are longer. The user is encouraged to create a multi-word phrase in the hope that the phrase's length will create so many possible character combinations that an intruder will be deterred from attempting a brute-force guessing attack. Even if a user

creates a passphrase about a meaningful detail of his/her life, it is still theoretically more secure than a detail-based password because of its greater length. Because it contains multiple words separated by spaces, computerized dictionary searches are ineffective against a passphrase.

4. Associative Passwords

Smith (1987) advocates a system which attempts to solve the password memorability problem by giving the user a "hint" about the password. Associative passwords employ a cue/response format. A user creates a list of cue words or short phrases and a response word/phrase to go with each cue. For instance, one of a user's cues might be "skiing"; the proper response might be "Keystone" (a ski resort in the Denver area). Smith suggests a profile of 20 cue/response pairs be created by each user. During log-on the user may be required to correctly respond to, say, five cues. Associative passwords theoretically offer a balanced mix of security and memorability. If the user avoids the use of easily guessed cue/response pairs (e.g., dog/cat, fast/slow, etc.), he/she can create a unique profile of password pairs that are resistant to intelligent guessing. Furthermore, associative passwords may be more easily remembered because the user is given the cue word/phrase to aid recall of the response word/phrase.

5. Cognitive Passwords

Cognitive passwords make use of the uniqueness of an individual's personal history, perceptions and opinions to confirm his/her identity. Initially, users are asked a set of simple questions, each of which seeks a short answer. The questions are styled so that the response from a particular person will be unique to that person. At the same time, the answer should not be common knowledge or publicly-available information. An example of a good cognitive password question might be *Who is your favorite professional entertainer?* The answer to such a question would obviously vary from person to person. Conversely, *On what date were you born?* is a poor question because the answer is publicly available (Zviran and Haga, 1990).

As is the case with associative passwords, a user initially creates a profile of cognitive passwords in response to a series of questions. During the log-on process, the user is required to correctly respond to one or more questions in order to be granted access. A properly-designed set of cognitive password questions will elicit a unique set of responses that are resistant to intelligent guessing. Furthermore, memorability should be improved because the user is required to simply remember the answer to an easy question he/she has answered before.

D. SUMMARY

Passwords are a widely employed method of computer user verification. Although the user-created password is the most well known and most frequently used, it is also prone to human frailties which often decrease its security. Other password formats exist which, in theory, offer the combination of increased security and greater ease of memorability.

III. RESEARCH METHODOLOGY

A. BACKGROUND

The principal goal of this study is to compare computer users' abilities to recall five password types over six intervals of time. Toward that goal, study participants were asked to create a series of passwords, then attempt to recall those passwords at a later date. There are two important differences between the methodology used in this study and that used in previous similar research (Beedenbender, 1990; Hulsey, 1989.) First, previous researchers collected their data through pencil-and-paper surveys, whereas this study employed a more realistic computer-based setting. Second, the recall abilities of participants in previous research were tested after only one interval three months. This study assigned each participant to one of six recall intervals: three days, one week, two weeks, one month, one-and-one-half months and two months. By collecting data at these different times, the study hoped to measure the decline in recall of passwords which would likely occur as the intervals lengthened. Additionally, there is the opportunity to compare the relative recall successes of the six password types with an eye toward determining if some types are more easily recalled at specific intervals.

B. METHODOLOGY

1. Data Collection Description

In order to simplify the gathering of data from six separate groups of participants (one for each of the six recall intervals) and to simulate an actual computer log-on environment, data collection was accomplished with the use of a computer program.

a. Study Participants

Graduate students in information systems management as well as in general management curricula participated in the study.

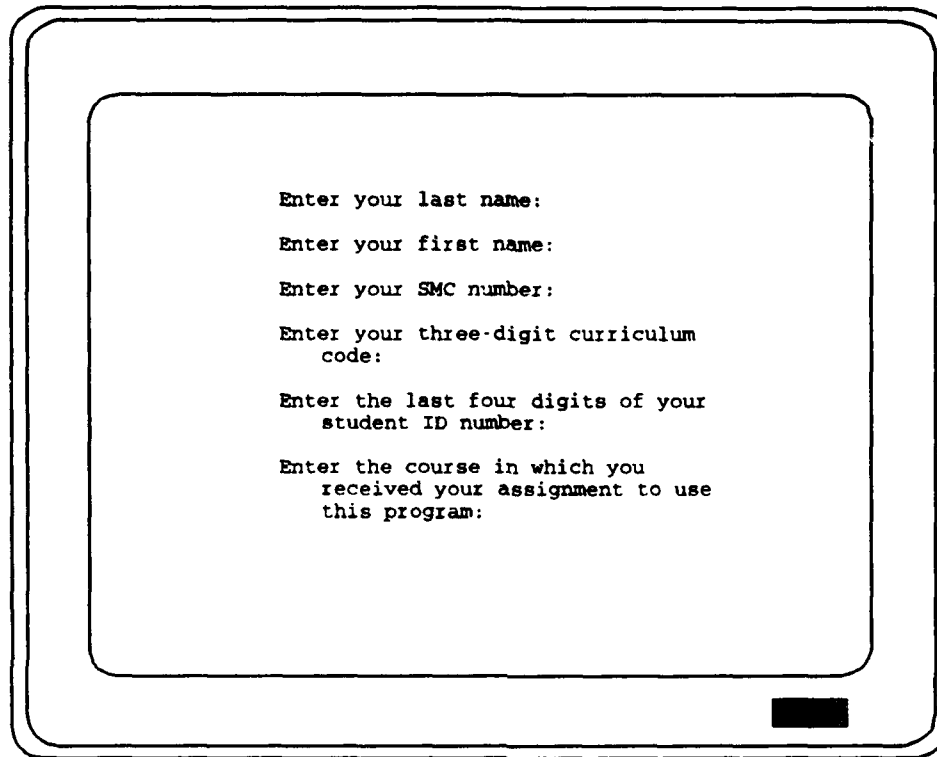
b. Detailed Study Description

Following printed instructions, study participants ran a simulation program installed on a local-area network in a microcomputer laboratory. The program provided each person with a basic understanding of the concepts being tested and his/her contribution to the research effort. The study introduction viewed by participants is shown in Appendix A.

c. Identifying Participants

Figure 1 illustrates personal identification data items collected from each participant during his/her first use of the program. Participants' names and student mailing center (SMC) box numbers were collected to allow reminder notices to be sent shortly before each participant's scheduled return visit. The three-digit curricular code indicates the type of degree a given student is pursuing. The last four digits of a student's ID number were used to create unique names for

computer files in which to store his/her password profile and recall data.



The diagram shows a computer monitor with a double-line border. Inside the monitor, there is a list of prompts for data collection, each followed by a blank line for input. The prompts are:

- Enter your last name:
- Enter your first name:
- Enter your SMC number:
- Enter your three-digit curriculum code:
- Enter the last four digits of your student ID number:
- Enter the course in which you received your assignment to use this program:

A small black rectangular box is located in the bottom right corner of the monitor's frame.

Figure 1 Identifying data collected from study participants

d. Creation of Password Data

After providing the computer program with the above information, a participant then viewed a series of five instructional screens, each of which assigned a password or asked him/her to create a password. Each study participant was assigned a computer-generated password in the first of these instructional screens (see Figure 2).

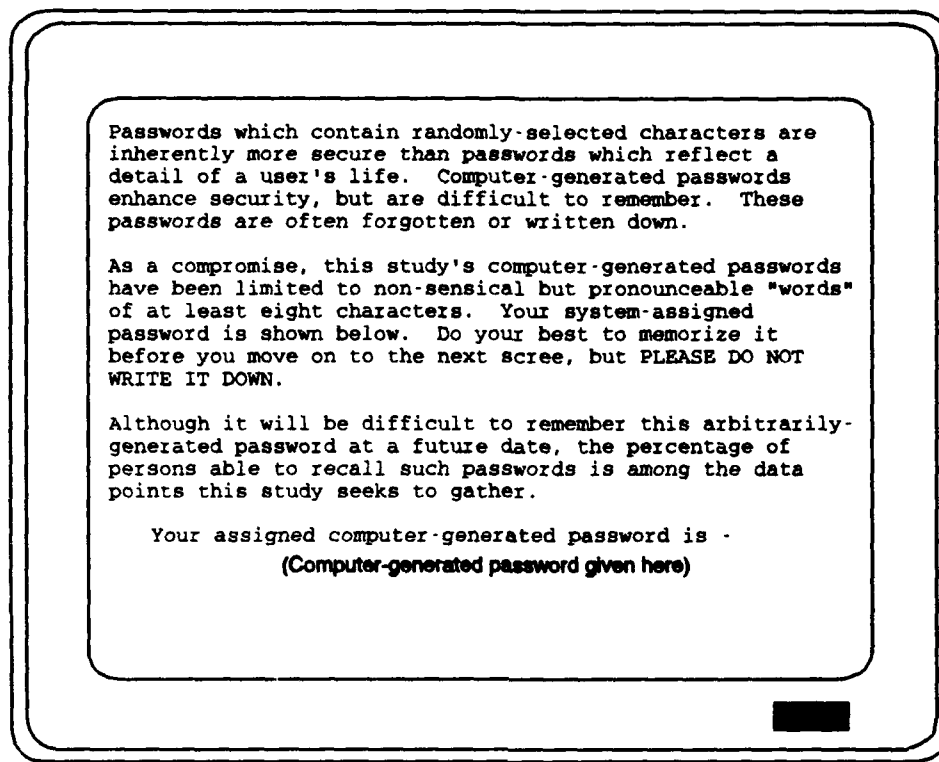


Figure 2 Assignment of computer-generated password

The next screens asked participants to create a single password of their own devising, a passphrase, 20 associative password combinations, and 20 cognitive passwords. The order in which these four screens were presented was randomized so that each of the 24 (4!) possible sequences was shown to an equal number of participants. This randomizing process was intended to guard against user fatigue during the session which might have caused passwords created in the later stages of the session to be less easily remembered. Since a similar number of participants created the four password types in any given order, the possibility of skewing overall recall results toward the

password types created in the early parts of the session was eliminated. The instructional screen and input mechanism for user-created passwords is shown in Figure 3.

User-created passwords are a commonly-used means of access control. In this section of the program, you will be asked to create a password. The password should be at least six and not more than ten characters long.

The password you create should contain no spaces; however, all other keyboard characters are permitted. Note that the computer considers uppercase and lowercase letters to be different. Thus, "ILuvMyCar" and "ILUVMYCAR" are NOT the same password.

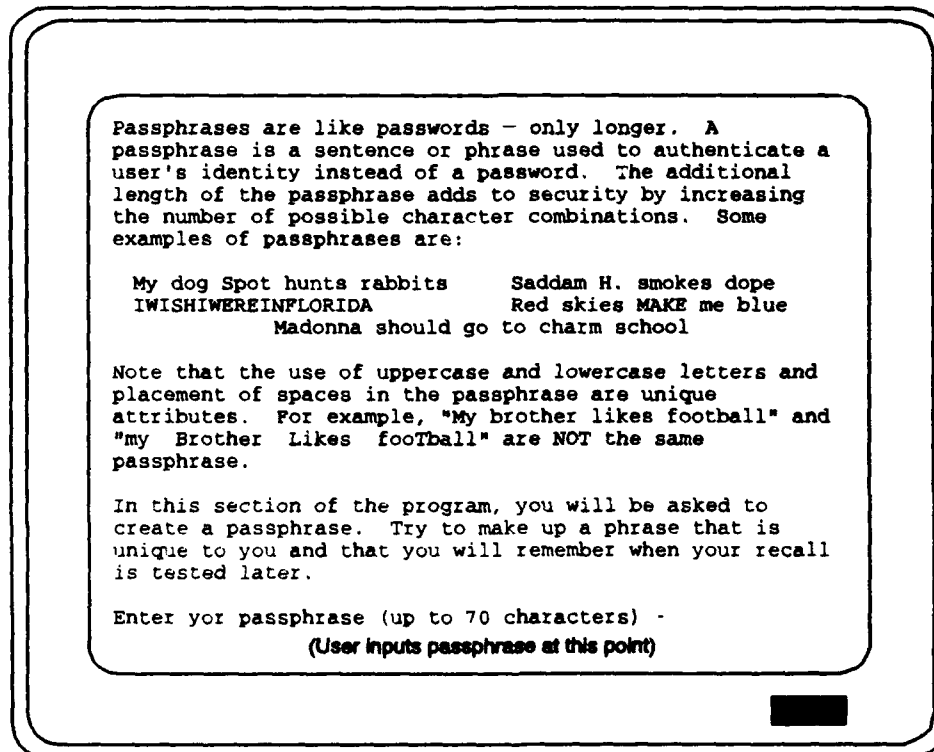
Do your best to make up a password which is unique to you and which you be able to recall when tested later.

Enter a password of your own choosing (6-10 characters) -
(User inputs his/her password at this point)

Figure 3 Instructions for the user-created password

Figure 4 depicts a computer screen describing the creation of a passphrase with attention drawn to the use of uppercase and lowercase letters and spaces. Because a mixture of cases can make a password or passphrase more secure, participants were allowed to mix them. The user-created passwords and passphrases devised by study participants were captured exactly as they were typed; an exact match was required in the recall phase to count as a successful simulated log-

on. An exact match of uppercase and lowercase letters was required for only these two password types. This case sensitivity was invoked in order to make the log-on simulation more closely reflect actual log-on practices. Inputs for system-created passwords and associative and cognitive passwords were not case-sensitive.



Passphrases are like passwords - only longer. A passphrase is a sentence or phrase used to authenticate a user's identity instead of a password. The additional length of the passphrase adds to security by increasing the number of possible character combinations. Some examples of passphrases are:

My dog Spot hunts rabbits	Saddam H. smokes dope
IWISHIWEREINFLORIDA	Red skies MAKE me blue
Madonna should go to charm school	

Note that the use of uppercase and lowercase letters and placement of spaces in the passphrase are unique attributes. For example, "My brother likes football" and "my Brother Likes football" are NOT the same passphrase.

In this section of the program, you will be asked to create a passphrase. Try to make up a phrase that is unique to you and that you will remember when your recall is tested later.

Enter your passphrase (up to 70 characters) -
(User inputs passphrase at this point)

Figure 4 Instructions for creation of a passphrase

Figure 5 shows the instructions that participants received to guide them through the creation of 20 sets of associative password cues and responses.

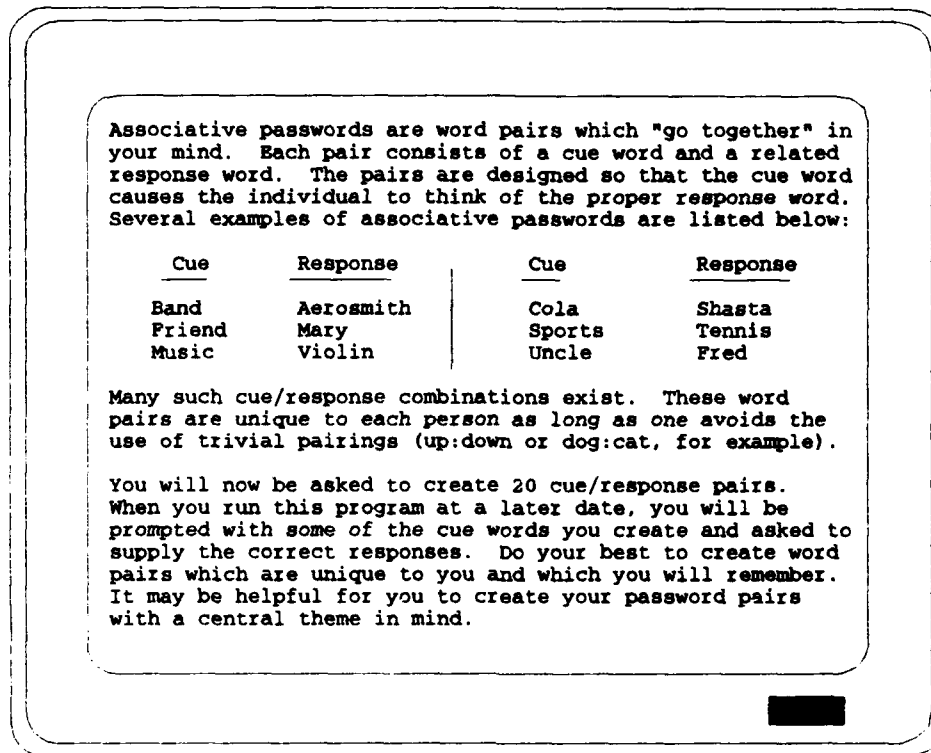


Figure 5 Instructions for creation of associative passwords

Participants were asked to create a profile of cognitive passwords. Figure 6 shows the instruction screen used to introduce participants to the cognitive password concept.

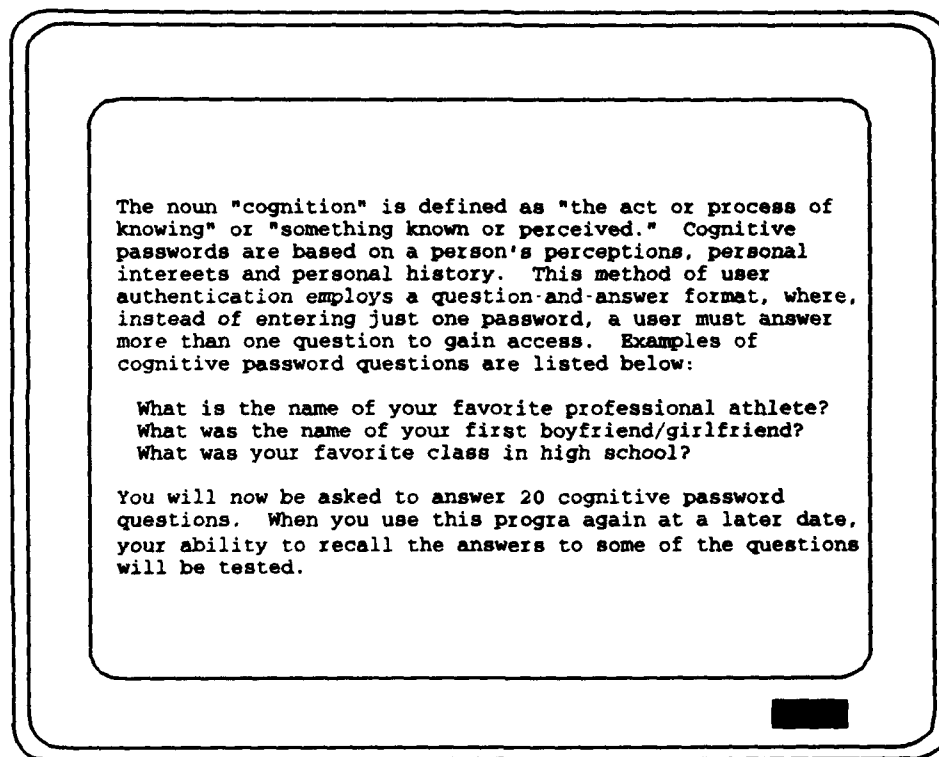


Figure 6 Instructions for the creation of cognitive passwords

Figure 7 shows the cognitive password questions used for this study. Some questions required objective answers which do not change (*From what elementary school did you graduate?*) while others ask for subjective opinions which may change over time (*What is your favorite restaurant?*).

e. Assignment of a Return Date

After creation of the five password types, each participant was assigned a time interval after which he/she was asked to return to the computer lab in order to recall the passwords in a simulated log-on session.

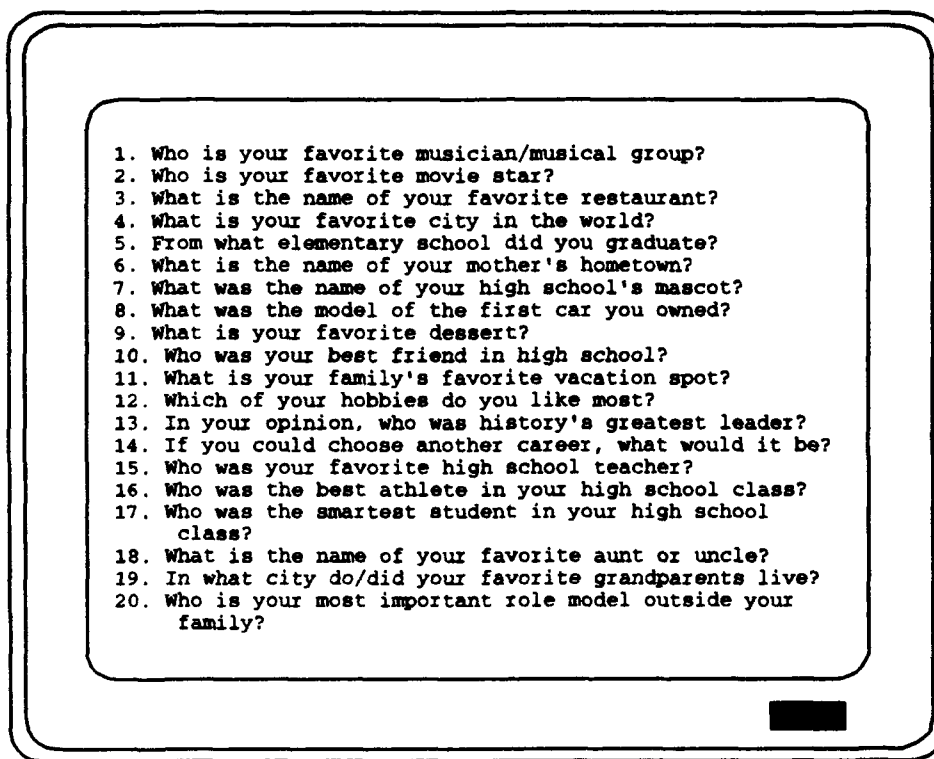


Figure 7 Cognitive password questions

Figure 8 shows the computer screen used to make this return date assignment. The computer program chose a return interval by cycling through the six intervals as persons used the program. Each succeeding participant was assigned the next interval. In this manner an even number of participants were assigned to each of the six intervals. Appendix B provides an overview of the recall intervals and their assignment.

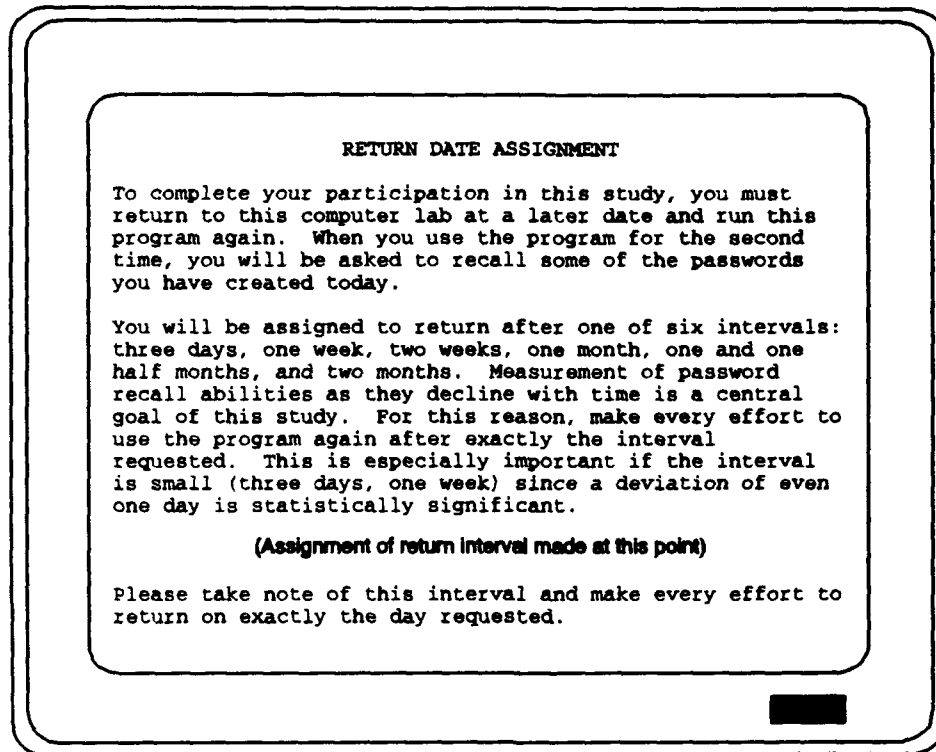


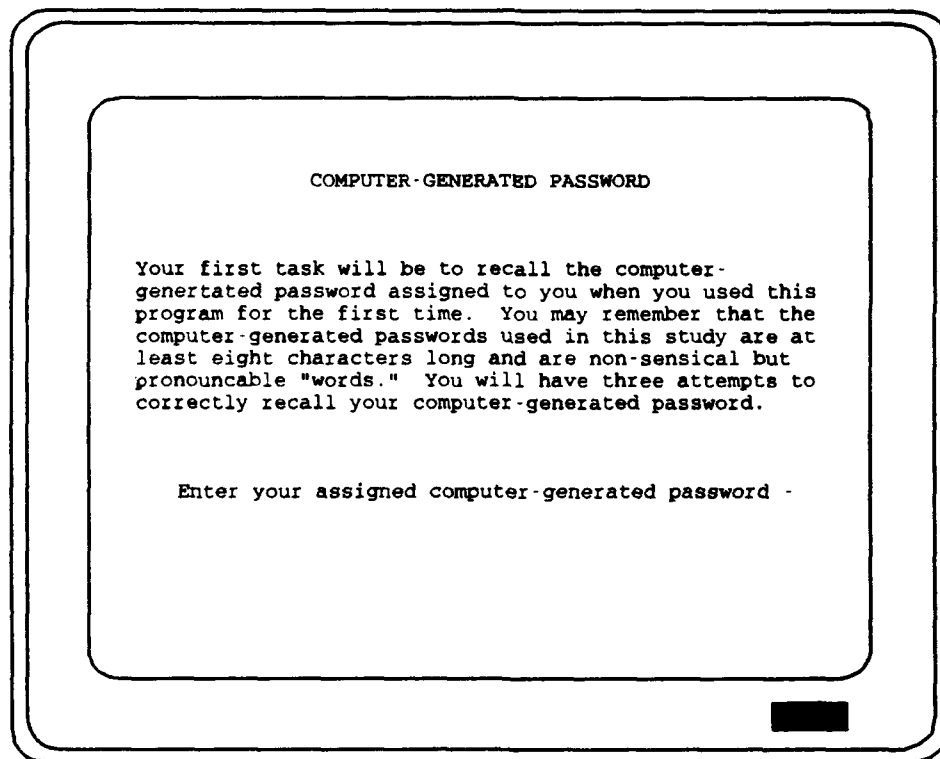
Figure 8 Assignment of a return date

2. Testing Recall of Passwords

Participants were reminded of the approach of their return date through notices placed in their student mailing center boxes using data on participant name and SMC number. At the return session, each participant ran the same computer simulation program he/she used to create passwords at the first session. Now, however, the program tested his/her recall of those passwords.

a. Testing Password Recall

As in the password creation phase of the study, participants were presented with an instruction screen to guide them through each step of recall testing. Figure 9 is the instruction screen used for recall testing of the participant's assigned computer-generated password. The datum collected by the computer program during this recall phase was the number of tries necessary for the user to successfully recall the password.



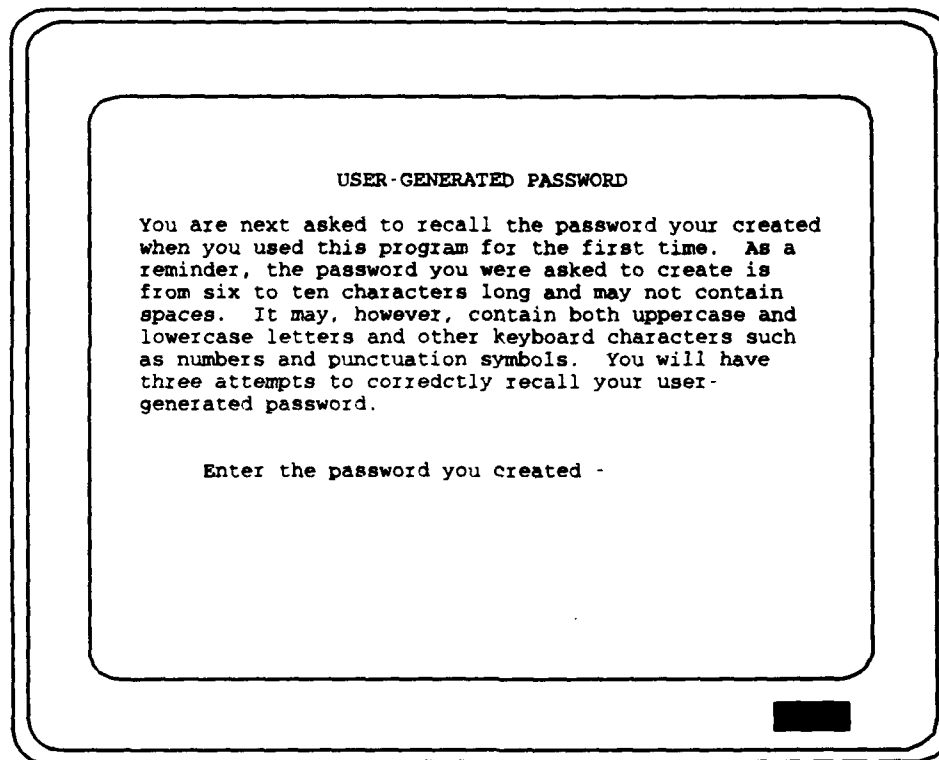
COMPUTER-GENERATED PASSWORD

Your first task will be to recall the computer-generated password assigned to you when you used this program for the first time. You may remember that the computer-generated passwords used in this study are at least eight characters long and are non-sensical but pronounceable "words." You will have three attempts to correctly recall your computer-generated password.

Enter your assigned computer-generated password - XXXXXXXXXX

Figure 9 Computer-generated password recall instructions

Figure 10 depicts the instruction screen presented to guide the participant through recall of his/her user-generated password. In the event that the participant had forgotten the distinctions between the various types of passwords, this and following screens served to refresh his/her memory. The instructions also contain a reminder about the possible use of both uppercase and lowercase letters in the chosen password.



USER-GENERATED PASSWORD

You are next asked to recall the password your created when you used this program for the first time. As a reminder, the password you were asked to create is from six to ten characters long and may not contain spaces. It may, however, contain both uppercase and lowercase letters and other keyboard characters such as numbers and punctuation symbols. You will have three attempts to correctly recall your user-generated password.

Enter the password you created -

Figure 10 User-generated password recall instructions

Figure 11 is the instruction screen used to prompt the user's recall of the passphrase he/she created. Again, the instructions remind the user about the characteristics of the password type being tested.

PASSPHRASE

You are next asked to recall the passphrase you created when you used this program for the first time. You may remember a passphrase is a phrase or short sentence of 70 characters or less. The placement of spaces and use of uppercase and lowercase letters within the phrase are unique attributes. You will have three attempts to correctly recall your passphrase.

Enter your passphrase - XXXXXXXXXX

Figure 11 Passphrase recall instructions

The simulated log-on using associative passwords required a user to correctly respond to five password cues in order for the log-on to be considered successful. From the 20 associative password pairs created by the participant, five cues were randomly selected by the computer. Feedback on the success of the each participant's recall was provided only after all five responses had been

given. In the event one or more of the responses were wrong, another set of five cues was randomly selected from the 20 pairs. Use of a cue-response pair in one group did not preclude it from being included in a subsequent group of five pairs. Figure 12 shows the instructions the user received for associative password recall.

ASSOCIATIVE PASSWORDS

Associative passwords, you may remember, are pairs of cue and response words or short phrases. You created twenty such pairs in your previous session with this program. The program will present you with five of the cues from your associative password data set; you will be asked to enter the proper response for each cue.

If your answer to one or more of the cues is incorrect, you will be asked to respond to a new set of five cues. No specific feedback about which responses are incorrect will be given; you will be told only that one or more responses are wrong. You will be allowed three attempts to supply a correct response to each of five cues.

Cue 1: Cue words supplied by	Response 1: User supplies
Cue 2: the computer program	Response 2: responses to each
Cue 3: using password data	Response 3: cue word.
Cue 4: files created by each	Response 4:
Cue 5: user.	Response 5:

Figure 12 Associative password recall information

An important facet of the test was the lack of feedback on which response a user might have gotten wrong. In the event of an error, only the fact that one or more of the responses was incorrect was reported to the user. Additionally, the associative password testing was non-case-sensitive. This contrasts with the case sensitivity

present in the user-created password and passphrase tests. While the use of mixed uppercase and lowercase letters was encouraged to create uniqueness among the latter two types of passwords, the associative password recall test did not include this feature. Therefore, the computer counted responses as correct regardless of their case.

Figure 13 is the instruction screen presented before recall testing of cognitive passwords.

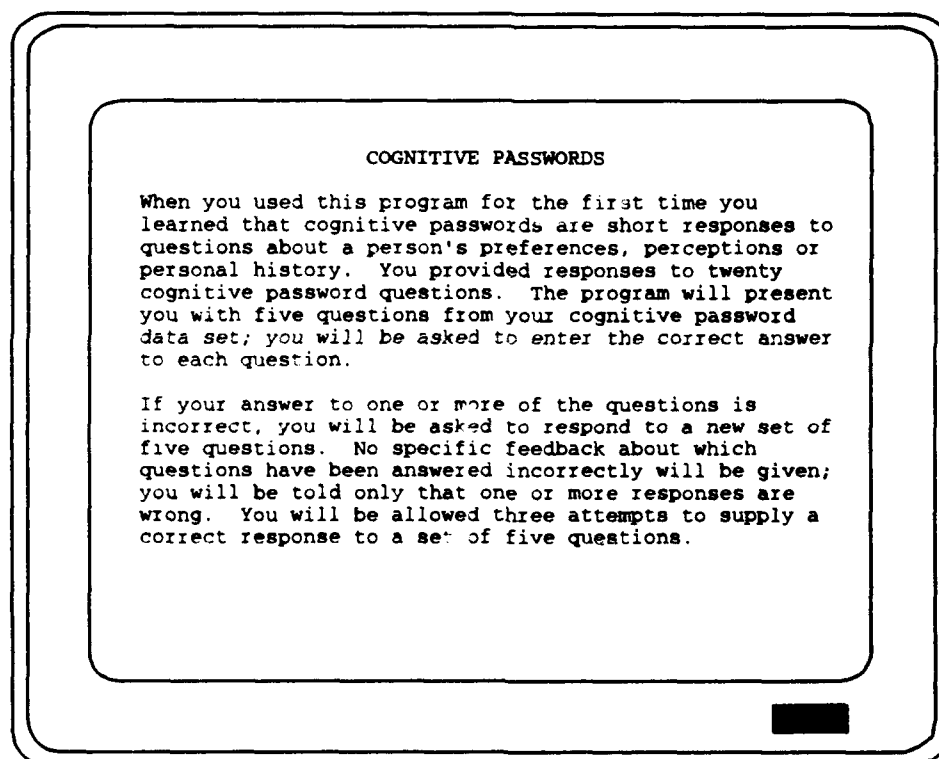


Figure 13 Cognitive password recall instructions

As was the case for associative recall testing, the study participant was given three chances to correctly answer five cognitive questions in a row; no specific feedback about errors was given. Since

the participant's memory of short answers to personal questions was the element being tested, the computer was not sensitive to the use of uppercase or lowercase letters.

b. Gathering Demographic Data

After the participant finished the five recall tests, his/her test scores were recorded in a computer file. He/she was then asked for several pieces of demographic information. Participants were questioned about their previous computer experience. Each was asked the number of years of previous computer experience he/she had and the types of computers (e.g., micro, mini, mainframe) he/she had used before. Each person was asked to rank the five password categories, first by ease of use then by ease of recall. Finally, participants were questioned about the mechanisms they used to remember their computer-generated password and the user password and passphrase they made up themselves. Each of these questions appeared only if the participant successfully recalled the type of password about which the question sought information. For instance, the user was asked how he/she remembered the passphrase only if he/she had correctly recalled it earlier in the program. Following these questions, the participant was presented a signoff screen thanking him/her for taking part in the study; the program then ended.

IV. FINDINGS

A. COMPARISON OF SIMULATED LOG-ON SUCCESSES

Of the 225 participants who were asked to take part in the password study, 183 completed their first session of the program. Of those, 164 returned to use the simulation a second time and complete the experiment; however, only participants who returned on or near the correct date (according to the interval assigned to them) or on a date corresponding to another of the study's six recall intervals provided usable data. Thus, 148 persons (66% of those asked to participate) contributed usable data to the study.

Table 1 provides a summary of recall successes as demonstrated by successful simulated log-ons.

TABLE 1
SUCCESSFUL SIMULATED LOG-ONS SUMMARIZED
BY PASSWORD TYPE

Recall Interval	No. of Persons	Computer Generated	User Created	Passphrase	Associative	Cognitive
3 days	24	13 (54%)	20 (83%)	18 (75%)	11 (46%)	13 (54%)
1 week	24	13 (54%)	16 (67%)	13 (54%)	11 (46%)	12 (50%)
2 weeks	29	20 (69%)	16 (55%)	9 (31%)	9 (31%)	13 (45%)
1 month	28	7 (25%)	8 (29%)	7 (25%)	8 (29%)	6 (21%)
1 months	20	2 (10%)	4 (20%)	2 (10%)	5 (25%)	2 (10%)
2 months	23	10 (43%)	7 (30%)	2 (9%)	4 (17%)	3 (13%)
Overall	148	65 (44%)	71 (48%)	51 (35%)	48 (32%)	49 (33%)

Figure 14 presents the recall data in graphic form.

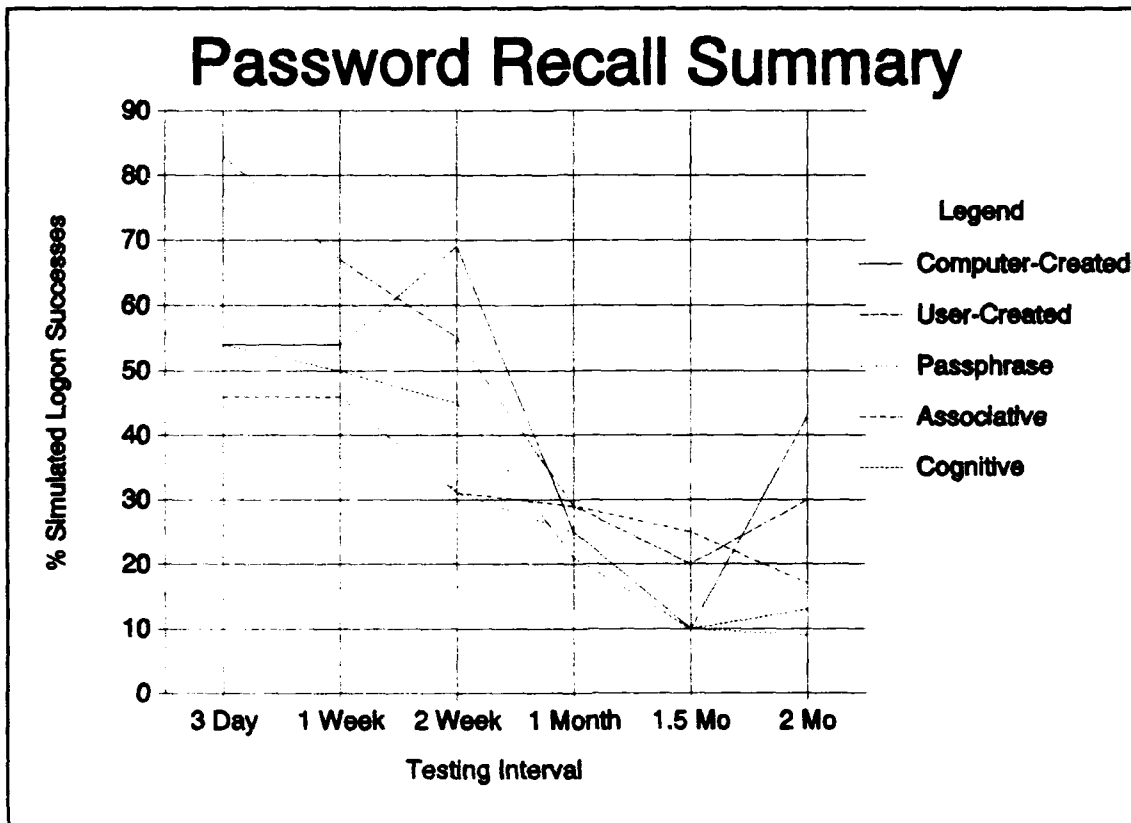


Figure 14

1. Recall of Associative and Cognitive Passwords

Because study participants were required to correctly answer five consecutive associative cues or cognitive questions to be credited with a successful log-on, the data in Table 1 and Figure 14 do not reflect the question-by-question success rate for these password categories. Table 2 provides a summary of the percentages of associative and cognitive passwords correctly remembered in each recall interval.

TABLE 2
RECALL SUCCESSES FOR ASSOCIATIVE
AND COGNITIVE PASSWORDS

Recall Interval	Associative	Cognitive
3 days	68%	75%
1 week	66%	74%
2 weeks	64%	70%
1 month	58%	58%
1½ months	45%	51%
2 months	51%	56%
All intervals	59%	63%

2. Recall success versus log-on attempts

For each of the five password categories, the study participant was permitted up to three log-on attempts. A correct response to any of the attempts constituted a successful simulated log-on. Table 3 provides a summary of the ability of all participants to correctly recall each password type on the first, second, or third try.

TABLE 3
SIMULATED LOG-ON SUCCESSES
BROKEN DOWN BY ATTEMPT

Password type	1st attempt successful	2nd attempt successful	3rd attempt successful
Computer-generated	36%	4%	3%
User-created	32%	11%	4%
Passphrase	29%	2%	3%
Associative	19%	8%	5%
Cognitive	18%	10%	5%

Data in Table 3 are a further breakdown of data in the final row of Table 1.

3. Recall of cognitive passwords

A summary of participants' abilities to answer each of the study's 20 cognitive password questions is presented in Table 4. The percentage in the table's second column is obtained by dividing the number of correct responses to that question by the total number of times the question was presented in the recall phase of the study.

TABLE 4
COGNITIVE PASSWORD RECALL
BROKEN DOWN BY QUESTION

Cognitive Password Question	Correct Responses
What is the name of your mother's hometown?	80%
From what elementary school did you graduate?	78%
In what city do/did your favorite grandparents live?	78%
Who is your favorite musician/musical group?	71%
Who was the smartest student in your high school class?	69%
What was the model of the first car you owned?	69%
What was the name of your high school's mascot?	68%
What is your favorite city in the world?	67%
Who was your best friend in high school?	67%
In your opinion, who was history's greatest leader?	67%
What is your favorite dessert?	66%
What is the name of your favorite aunt or uncle?	66%
Who is your favorite movie star?	60%
Where is your family's favorite vacation spot?	60%
If you could choose another career, what would it be?	60%
Who was your favorite high school teacher?	59%
What is the name of your favorite restaurant?	58%
Who was the best athlete in your high school class?	56%
Which of your hobbies do you like most?	55%
Who is your most important role model outside your family?	38%

B. METHODS OF RECALL

1. Computer-generated passwords

Each participant who was able to correctly recall his/her computer-generated password, user-generated password and passphrase was asked to specify the recall mechanism he/she used. Table 5 summarizes the recall methods used by participants who correctly remembered the computer-generated passwords assigned them.

TABLE 5
METHODS USED TO RECALL
COMPUTER-GENERATED PASSWORDS

Method of Recall Used	Participants using this method
Wrote it down	1 (2%)
Remembered because it was pronounceable	32 (49%)
No special method used	9 (14%)
Other	23 (35%)
Total successful log-ons	65

As shown in the table, the study's deliberate creation of non-sensical but pronounceable words had an effect on the number of persons able to recall the assigned passwords. Just under half of those who remembered their computer-generated passwords were able to do so because the "word" was pronounceable. Association of the

assigned word with some phrase or detail was the second most popular method for jogging the memory of participants. Of those who specified "Other" for their recall method, 20 of 23 said they used a word association scheme to help them remember the assigned password.

2. User-created passwords

The category most remembered by the study's participants was the user-created password. Participants seemed especially able to recall this password type during the two smallest recall intervals; successful log-ons with user-created passwords outnumbered other password categories for the three-day and one-week intervals. Table 6 summarizes the recall methods used by those participants who successfully logged on with their user-created password.

TABLE 6
METHODS USED TO RECALL USER-CREATED PASSWORDS

Method of Recall Used	Participants using this method
Wrote it down	0
Password I've used before	21 (30%)
Significant detail in my life (date, name, etc)	26 (37%)
Invented a pronounceable word	8 (11%)
No special method used	6 (9%)
Other	10 (14%)
Total successful log-ons	71

3. Passphrases

Those participants who correctly recalled their passphrase were asked to indicate how they did so. Table 7 presents a summary of the recall methods used.

TABLE 7
METHODS USED TO RECALL PASSPHRASES

Method of Recall Used	Participants using this method
Wrote it down	0
Adapted it from a password	1 (2%)
Significant detail in my life (date, name, etc)	14 (28%)
A phrase I use or hear frequently	18 (35%)
No special method used	9 (18%)
Other	9 (18%)
Total successful log-ons	51

C. EASE OF PASSWORD RECALL

Study participants were asked to rank the five password types in order of the ease with which each could be recalled. Table 8 provides a summary of participants' responses. The number of participants who ranked each password type first through fifth (easiest through most difficult) for ease of recall is noted. The mean score for each row of the table is computed by multiplying the number in each column by the ranking that column represents, summing the five products, then dividing the sum by the number of persons who responded to this

question (143). Since a greater ease of recall is indicated by a low numerical ranking, the password category with the lowest mean score is the one the study's participants collectively judged easiest to remember.

TABLE 8
RANKING OF PASSWORD MECHANISMS
ACCORDING TO EASE OF RECALL

Password Type	Ease of Recall Ranking (number of persons)					Rank	Mean Score
	1st	2nd	3rd	4th	5th		
User-created	56	25	28	27	10	1	2.43
Associative	23	38	38	30	13	2	2.78
Cognitive	22	33	27	38	23	3	3.05
Passphrase	16	28	43	35	20	4	3.08
Computer-generated	26	19	7	13	77	5	3.65

The above ranking of the five password types according to ease of recall agrees exactly with previous research (Beedenbender, 1990), although mean scores were more tightly bunched in this study.

D. EASE OF PASSWORD USE

Study participants were asked to rank each of the five password categories according to which of them was easiest to use. This question was posed as an attempt to remove recall criteria from the ranking process. Participants were specifically told to assume they recalled each password type equally well and to rank them on the

basis of ease of use only. Table 9 is organized in the same manner as Table 8; mean scores and rankings are determined in the same way. The number of persons who responded to this question was 147. For ease of use, user-created passwords and associative passwords are once again the preferred password methods. The last three password types have mean scores that are very close together, indicating ambivalence on the part of participants when asked to choose between them.

TABLE 9
RANKING OF PASSWORD MECHANISMS
ACCORDING TO EASE OF USE

Password Type	Ease of Use Ranking (number of persons)					Rank	Mean Score
	1st	2nd	3rd	4th	5th		
User-created	76	33	15	22	1	1	1.90
Associative	19	36	31	39	23	2	3.10
Computer-generated	22	32	21	16	55	3	3.30
Cognitive	16	20	33	42	33	4	3.33
Passphrase	12	26	48	28	35	5	3.37

The results in Table 9 closely match previous research (Beedenbender, 1990). The first and second rankings are identical, while the latter three rankings are not ordered the same; however, the small difference in preferences between the final three categories probably makes this observation insignificant.

E. COMPUTER EXPERIENCE OF PARTICIPANTS

Since this study's participants were students in a graduate curriculum, each of them had previously earned a bachelor's degree and most had worked with computer equipment during at least part of their careers.

1. Length of prior computer experience

Participants were asked to indicate the number of years of computer experience they had before taking this study. For the purposes of the study, computer experience was defined as formal computer education or regular use of a computer at work or at home. Table 10 summarizes the study participant's experience levels. For all 148 participants, the average number of years of computer experience was 5.0; the median number was 4. Eight of the study's participants said they had no previous computer experience; one person had as many as 19 years of experience.

TABLE 10
STUDY PARTICIPANTS' COMPUTER EXPERIENCE

Previous Computer Experience	Number of Participants
≤ 1 year	31
> 1 year and ≤ 3 years	35
> 3 years and ≤ 5 years	34
> 5 years and ≤ 7 years	11
> 7 years and ≤ 9 years	11
> 9 years	26

2. Types of computers used

The variety of computer work done by each participant was cataloged further by type of computer. All but two of the participants said they had used a microcomputer (personal computer) before. Following microcomputers, the next most used architecture was the mainframe computer. Table 11 gives a summary of this data.

TABLE 11
STUDY PARTICIPANTS' EXPERIENCE
WITH COMPUTER ARCHITECTURES

Computer Architecture	Persons with experience
Microcomputer	146
Microcomputer network	78
Minicomputer workstation	45
Microcomputer with modem	69
Mainframe computer terminal	88

V. DISCUSSION

A. ANALYSIS OF SUCCESSFUL LOG-ON RATES

Raw data presented in Table 1 and their graphical depiction in Figure 14 show a declining rate of successful simulated log-ons over the course of the six recall intervals. These results are intuitively plausible: decreasing success in remembering passwords would be expected as time between the first and second computer sessions is increased. A statistical analysis provides a more quantitative examination of any observed differences in log-on successes.

1. Description of Analysis

In order to examine the data for differences of recall rates between the various password types at each recall interval, a chi-square goodness-of-fit test was employed. This test is appropriate for random, independent samples in which the observations being tested fall into only one of a series of mutually exclusive and collectively exhaustive categories (Porter and Hamm, 1986, pp. 183-193). In this study, each of the five password tests evaluates to one and only one of two possible results: successful log-on or unsuccessful log-on. Since the computer program sessions were conducted by each person individually, his/her test results are independent of any other person's. The null hypothesis (H_0) and alternative hypothesis (H_1) used for the goodness-of-fit test are listed below.

H_0 : There are no significant differences between successful simulated log-on rates for the five password categories

H_1 : There are significant differences between successful simulated log-on rates for the five password categories

Tests were performed on data for each of the six recall intervals and on all interval results collectively. A .05 level of significance ($\alpha = .05$) was used as the accept/reject criterion.

2. Results of Analysis

Table 12 summarizes the results of the seven goodness-of-fit tests performed. Detailed results of each test are presented in Appendix C.

TABLE 12
RESULTS OF CHI-SQUARE TESTS
OF DIFFERENCES IN PASSWORD RECALL RATES

Recall Interval	No. of Participants	Accept/Reject H_0 at $\alpha=.05$
3 days	24	Reject
1 week	24	Accept
2 weeks	29	Reject
1 month	28	Accept
1½ months	20	Accept
2 months	23	Reject
Overall	148	Reject

The null hypothesis is rejected for the three-day, two-week, and two-month recall intervals as well as for the overall data set.

The results of the six interval tests agree with an intuitive analysis of Figure 14. Simulated recall results for the five password types are clustered together at the one-week, one-month and one-and-one-half-month testing intervals, indicating a similarity in recall rates for each category. On the other hand, data at the three-day, two-week and two-month intervals are spread across wider ranges of values, suggesting there are statistically significant differences in the recall rates.

3. Simulated Log-on Versus Individual Password Recall

Table 2 reported a question-by-question recall success rate for associative and cognitive passwords. A comparison of Table 1 and Table 2 data reveals that, while all-interval simulated log-on successes for associative and cognitive passwords were 32% and 33%, respectively, recall rates for those passwords were much higher, 59% and 63%, when results are tabulated on a question-by-question basis. As would be expected, the success rate when the questions are considered one at a time is much greater than when five in a row must be correctly answered. This has implications for computer managers who might use an associative or cognitive authentication scheme to grant user access. These data suggest that altering the conditions which define a log-on success would result in a decrease in the rejection of bona fide system users. Since individual question recall

was near 60% for both associative and cognitive passwords, requiring a user to correctly answer only three of five associative cues or cognitive questions would achieve a greater log-on success rate. The 60% figure likely would not degrade security since these data show approximately 40% of associative and cognitive questions are answered incorrectly due to forgetfulness. As an alternative to lowering the required recall rate, lowering the number of questions required for log-on would likely also increase log-on success rates.

4. Benefits of Permitting Multiple Log-on Attempts

Table 3 shows substantial differences between recall successes on first attempts and follow-on attempts for computer-generated and user-created passwords and passphrases. When comparing the success rates for associative and cognitive passwords, the differences are not as great. In fact, the number of successful simulated log-ons on second and third attempts combined is near that achieved on the first attempt in the associative and cognitive categories. The reason for this is clear when the study's log-on requirements are reviewed. A participant who fails on his/her first log-on attempt by incorrectly responding to one or more of five questions is asked another five questions randomly from the pool of 20 associative cue-response pairs or cognitive answers each participant provided. The respondent is thus given five new questions to answer. This contrasts with the computer-generated password, user-created password and passphrase log-on schemes where the participant is given three opportunities to

correctly recall the same password or passphrase. A person who has failed once to remember a password/passphrase is less likely to recall that same word/phrase given another attempt than a person provided with a set of associative or cognitive questions that *differ* from those asked previously. Given this, one might assert that associative and cognitive simulated log-on success rates for each attempt *should be equal*. The decline in success rates for each subsequent attempt in these categories might be explained by either or both of two possibilities. First, a previously-asked question which the participant answered incorrectly might appear again in a later log-on attempt (five questions are chosen at random from the entire question database for each log-on attempt). Second, a participant might become discouraged by failure in his/her first attempt and lose interest in follow-on attempts.

B. ANALYSIS OF RECALL MECHANISMS

1. Significance of Cognitive Password Recall

While reviewing Table 4, note that most of the questions whose recall rates were the highest require objective answers which do not change over time (the answers to these questions are established facts). There were four such questions in this study: *From what elementary school did you graduate?*, *What is the name of your mother's hometown?*, *What was the name of your high school's mascot?* and *What was the model of the first car you owned?* Two of

these questions were first and second in recall rate; the other two are in the top six. Questions in the lower part of the chart ask for more subjective responses which may change with time or the whims of the participant. The observed ability of study participants to more easily recall objective cognitive password questions agrees with the results of previous research (Beedenbender, 1990; Hulsey, 1989). The implication for the administrator of a computer system which uses cognitive passwords is clear: deliberately designing the cognitive password questions so that they require objective vice subjective answers will increase the authorized user's password recall rate, thus reducing rejections of authorized users.

2. How Secure Are Our Passwords?

Regarding data in Table 6, an implication about the security of user-created passwords lies in the observation that nearly 70% of participants who recalled their passwords did so because the passwords were re-used or represented a significant detail of their lives. The regular changing of passwords and avoidance of passwords containing publicly available personal information (phone number, anniversary date, Social Security Number, etc) are tenets of good password security. It appears that many of the study's participants either were not aware of or ignored these concepts.

Conclusions about the probable security awareness of participants who recalled the passphrase they created may be drawn from the data in Table 7. Over 60% of those who recalled their

passphrase were able to do so because it was related to a significant detail in their life or a phrase they or someone they knew used frequently. Although its length makes it more secure than a password, basing a passphrase on personally-related data may make the phrase vulnerable to intelligent guessing by outsiders.

Previous research supports the conclusion that computer users may not practice good security when they create their own passwords. Beedenbender (1990), Sawyer (1990) and Hulsey (1989) found that 77%, 78% and 78%, respectively, of those surveyed used a meaningful detail or combination of meaningful details about their lives to create their password. While this enables the user to more easily remember his/her password, users must be careful to avoid building an easily guessable password when they incorporate details of their lives into password creation.

C. PERCEPTIONS VERSUS RECALL RESULTS

A comparison of data displayed in Table 1, Table 8, and Table 9 reveals that participants' feelings about the ease of use and ease of recall of a given password type were not necessarily related to the simulated log-on success they experienced for that type. Note that, with an overall recall rate of 44%, computer-generated passwords were the second most frequently recalled password (Table 1). Despite this, Table 8 data show computer-generated passwords were subjectively rated the least easy to remember. When only participants' most-

easily-remembered rankings are considered, computer-generated passwords (chosen by 26 persons) are judged the second most easily recalled type. This high ranking is more than offset, however, by the large number of participants (77) who ranked computer-generated passwords the most difficult to remember. Table 16 provides a summary.

TABLE 16
COMPARISON OF LOG-ON SUCCESS
WITH SUBJECTIVE RANKINGS

Ranking	Successful Simulated Log-ons	Ease of Recall	Ease of Use
1	User-created	User-created	User-created
2	Computer-generated	Associative	Associative
3	Passphrase	Cognitive	Computer-generated
4	Cognitive	Passphrase	Cognitive
5	Associative	Computer-generated	Passphrase

The Table 16 summary shows the user-generated password was the most frequently recalled and also the most preferred from an ease of recall and ease of use standpoint. Although associative passwords ranked second in both subjective evaluation categories, study participants were able to successfully log-on using associative passwords less frequently than any other password type. Insight into

possible reasons for this seeming discrepancy can be gained by remembering the data in Table 2 (discussed in paragraph A.3. above). When the recall rates are defined question-by-question vice by log-on successes, associative passwords are correctly remembered more frequently than any other category except cognitive passwords. It might be inferred that participants subjectively ranked cognitive and associative passwords more highly because they judged them on a question-by-question basis instead of on the basis of log-on success.

VI. CONCLUSIONS

A. THE "BEST" PASSWORD TYPE

The principal goal of this study was to determine which, if any, of the five password types tested could be consistently remembered better than the others. The study measured password recall by the yardstick of simulated computer log-ons, just as would occur in the real world. Test results summarized in Table 12 show there was no consistent significant difference in the log-on rates of the different password types. An identical conclusion can be reached through examination of Table 1 and Figure 14. The recall rankings of the five passwords shift for every set of intervals. There is no clear overall "winner" with respect to memorability. If a given type must be declared the most consistently remembered, it is the user-created password, which held or shared the highest log-on success rate in three of the six recall intervals.

B. PAPER SURVEYS VERSUS COMPUTER STUDIES

The above conclusion does not agree with previous research. Beedenbender (1990) found graduate students were able to remember cognitive and associative passwords at rates two to three times that of computer-generated passwords, user-created passwords or passphrases. Hulsey (1989) obtained results similar to Beedenbender's

when comparing recall rates of cognitive passwords with those of computer-generated and user-created passwords (Hulsey did not test passphrases or associative passwords). The failure of cognitive and associative passwords to outscore other password categories in the study presented here is almost certainly due to the differing conditions under which this study was conducted.

1. Comparison of Survey Methods

Hulsey and Beedenbender administered their surveys using a paper format. The graduate students who made up their study groups created password profiles by filling out questionnaires. After three months, the participants were asked to try to recall their passwords. In the case of associative and cognitive passwords, participants were presented with the entire set of associative cues and cognitive questions at once and asked to respond to them. In contrast, participants in this study interacted with a microcomputer both during the assignment/creation of passwords portion of the study and during the recall portion of the study.

2. Comparison with Previous Results

The randomly selected five associative password cues and five cognitive password questions provide less of a jog to the memory than seeing all the questions at once. This is likely the reason that recall averages for associative and cognitive questions (from Table 2) were 59% and 63%, ten percent less than the 69% and 74% achieved by participants in the Beedenbender study. The difference is even

greater when compared with Hulsey's results. His participants successfully recalled 82% of their cognitive passwords after three months. This study's lower recall averages occurred even though every recall interval was shorter than the three month interval applied to all of Beedenbender's and Hulsey's participants.

More important than the comparison of overall recall averages is the relationship between successful simulated log-ons this study measured. When the study's criterion of log-on completion is applied as the metric of success, users found success with computer-generated and user-created passwords 10%-15% more frequently than with associative or cognitive passwords, a result completely opposite from previous research.

3. Theory Versus Practice

The bottom-line conclusion one must reach from these observations is that, while graduate students tested with paper questionnaires were able to recall associative and cognitive passwords markedly better than other password types, graduate students required to complete a simulated computer log-on found these two methods the least successful. The difference exposed here is the difference between success of a concept in theory and in practice. In the closer-to-the-real-world conditions under which this computer-based study was conducted, conclusions reached by Beedenbender and Hulsey that associative or cognitive passwords are a better means of user authentication than more traditional password systems cannot be

supported. As noted in section VI.A. above, data from this study produce no clear winner. Perhaps future studies will clarify the inconsistencies.

APPENDIX A

The introduction and instructions below were presented to study participants during their first session with the computer program.

Most computer systems which employ access controls authenticate a user's identity with a password scheme. A user who supplies the correct password is allowed access to the computer's resources. Each password is theoretically unique to each authorized computer user. In reality, however, password methods have varying levels of security effectiveness. Those methods which are easiest to use are often the least secure.

An individual may employ his/her telephone number as a password to gain access to a computer; that person would likely have little trouble remembering (and thus using) the password. Unfortunately, a person's telephone number, even if unlisted, is available to many people. An intruder with a small amount of resourcefulness might gain unauthorized access to the protected computer by simply trying such an obvious possibility. This "intelligent guessing" of a person's birthday, anniversary, Social Security number, spouse's name or other common knowledge is a leading means used to foil computer security mechanisms.

This computer program and your use of it are part of a study to compare the ease of use and security of five methods of user authentication. As the program progresses, you will be assigned a computer-generated password and asked to create a user-generated password, a passphrase, a profile of associative passwords and a profile of cognitive passwords. Each of these terms will be explained as the program continues.

At the end of the program, you will be given a future date on which you are to return to this computer lab and run this program again. When you use the program the second time, your recall of the five types of passwords will be tested. Results of students' recall tests will be tabulated and compared to determine which password

mechanism offers the best combination of security and ease of use.

Some final words: please do not make any notes about the password data you provide today. The ability of each person to recall his/her passwords without the help of written notes is the most important quantity this study seeks to measure.

APPENDIX B

Table B-1 below provides an overview of the password study's organization. During their first computer session, participants were assigned membership in one of six password recall intervals. Each row of the table corresponds to an interval. The data entry phase of the study, during which participants created their passwords, is noted by a D. The recall (observation) phase of the study occurs at the interval noted with an O. Subscripts indicate the recall interval to which the letter applies.

TABLE B-1
ORGANIZATION OF PASSWORD
STUDY RECALL INTERVALS

Recall Group	First Session	Second Session					
		3 day	1 wk	2 wk	1 mo	1½ mo	2 mo
1	D ₁	O ₁					
2	D ₂		O ₂				
3	D ₃			O ₃			
4	D ₄				O ₄		
5	D ₅					O ₅	
6	D ₆						O ₆

APPENDIX C

Displayed below are the results of the seven chi-square goodness-of-fit tests performed on the simulated log-on data collected by this study. The goodness-of-fit calculations were performed using *The Student Edition of MINITAB* (Release 1.1).

Tests were performed on data for each of the six recall intervals and on all interval results collectively. A .05 level of significance ($\alpha=.05$) was used as the accept/reject criterion. Each of the tests involved five password categories. Since the number of degrees of freedom for a chi-square goodness-of-fit test is simply one less than the number of categories of observations, four degrees of freedom ($df=4$) were present in each test. In order to reject the null hypothesis in a given test, the test's chi-square statistic must be greater than or equal to 9.49, which is the value of χ^2 for $\alpha=.05$ and $df=4$ (Porter and Hamm, 1989, p. 394).

The goodness-of-fit test for an equally likely model (in which the likelihood of success or failure for each category is equal) arrives at its chi-square test value by comparing the observed number in each category with the expected value of each category. The chi-square test statistic is computed through use of the formula

$$\chi_o^2 = \sum \frac{(O-E)^2}{E}$$

where

χ_o^2 = the chi-square statistic computed from the sample data
(the "O" subscript refers to the "observed" statistic)

O = the observed value in each category

E = the expected value in each category

Chi-Square Test for Three-Day Interval

	CGPW	UCPW	PPHR	ASPW	COPW	Total
Log-on	13	20	18	11	13	75
Successes	15.00	15.00	15.00	15.00	15.00	
Log-on	11	4	6	13	11	45
Failures	9.00	9.00	9.00	9.00	9.00	
Total	24	24	24	24	24	120

$$\text{ChiSq} = 0.267 + 1.667 + 0.600 + 1.067 + 0.267 + 0.444 + 2.778 + 1.000 + 1.778 + 0.444 = \underline{10.311}$$

$$\text{df} = 4$$

Chi-Square Test for One-Week Interval

	CGPW	UCPW	PPHR	ASPW	COPW	Total
Log-on	13	16	13	11	12	65
Successes	13.00	13.00	13.00	13.00	13.00	
Log-on	11	8	11	13	12	55
Failures	11.00	11.00	11.00	11.00	11.00	
Total	24	24	24	24	24	120

$$\text{ChiSq} = 0.000 + 0.692 + 0.000 + 0.308 + 0.077 + 0.000 + 0.818 + 0.000 + 0.364 + 0.091 = \underline{2.350}$$

$$\text{df} = 4$$

Chi-Square Test for Two-Week Interval

	CGPW	UCPW	PPHR	ASPW	COPW	Total
Log-on	20	16	9	9	13	67
Successes	13.40	13.40	13.40	13.40	13.40	
Log-on	9	13	20	20	16	78
Failures	15.60	15.60	15.60	15.60	15.60	
Total	29	29	29	29	29	145

$$\text{ChiSq} = 3.251 + 0.504 + 1.445 + 1.445 + 0.012 + 2.792 + 0.433 + 1.241 + 1.241 + 0.010 = \underline{12.375}$$

$$\text{df} = 4$$

Chi-Square Test for One-Month Interval

	CGPW	UCPW	PPHR	ASPW	COPW	Total
Log-on	7	8	7	8	6	36
Successes	7.20	7.20	7.20	7.20	7.20	
Log-on	21	20	21	20	22	104
Failures	20.80	20.80	20.80	20.80	20.80	
Total	28	28	28	28	28	140

$$\text{ChiSq} = 0.006 + 0.089 + 0.006 + 0.089 + 0.200 + 0.002 + 0.031 + 0.002 + 0.031 + 0.069 = \underline{0.524}$$

$$\text{df} = 4$$

Chi-Square Test Results for One-and-one-half-month interval

	CGPW	UCPW	PPHR	ASPW	COPW	Total
Log-on	2	4	2	5	2	15
Successes	3.00	3.00	3.00	3.00	3.00	
Log-on	18	16	18	15	18	85
Failures	17.00	17.00	17.00	17.00	17.00	
Total	20	20	20	20	20	100

$$\text{ChiSq} = 0.333 + 0.333 + 0.333 + 1.333 + 0.333 + 0.059 + 0.059 + 0.059 + 0.235 + 0.059 = \underline{3.137}$$

$$\text{df} = 4$$

The presence of five cells with expected counts less than 5.0 indicates the chi-square test statistic probably is not accurate. This set of data does not lend itself toward goodness-of-fit testing.

Chi-Square Test for Two-Month Interval

	CGPW	UCPW	PPHR	ASPW	COPW	Total
Log-on	10	7	2	4	3	26
Successes	5.20	5.20	5.20	5.20	5.20	
Log-on	13	16	21	19	20	89
Failures	17.80	17.80	17.80	17.80	17.80	
Total	23	23	23	23	23	115

$$\text{ChiSq} = 4.431 + 0.623 + 1.969 + 0.277 + 0.931 + 1.294 + 0.182 + 0.575 + 0.081 + 0.272 = \underline{10.635}$$

$$\text{df} = 4$$

Legend:

CGPW - Computer-Generated Password
UCPW - User-Created Password
PPHR - Passphrase
ASPW - Associative Password
COPW - Cognitive Password

Expected values are listed below observed values for each password category and simulated log-on outcome (success/failure).

LIST OF REFERENCES

Alexander, Michael, *Biometric System Use Widening*, Computerworld, 24:2, Jan 8, 1990, pp. 15-16.

Alexander, Michael, *Poor Security Made DoD Easy Hacker Prey*, Computerworld, 25:47, Nov 25, 1991, p. 92.

Beedenbender, Mark G., *A Comparison of Password Techniques*, Master's Thesis – Naval Postgraduate School, Mar 1990.

Fitzgerald, Karen, *The Quest for Intruder-Proof Computer Systems*, IEEE Spectrum, 26:8, Aug 1989, pp. 22-26.

Hulsey, John Douglas, *Cognitive Passwords: The Key for Effective Access Control*, Master's Thesis – Naval Postgraduate School, Sep 1989.

Joyce, Rick and Gupta, Gopal, *Identity Authentication Based on Keystroke Latencies*, Communications of the ACM, 33:2, Feb 1990, pp. 168-176.

Mital, D.P. and Lau, K.T., *A Microprocessor-Based Signature Verification System*, IEEE Transactions on Consumer Electronics, 35:4, Nov 1989, pp. 849-851.

Padovano, Michael, *Five Ways to Foil Password "Crackers"*, Systems Integration, 24:9, Sep 1991, p. 29.

Parks, J.R., *Personal Identification – Biometrics*, Proceedings of the Seventh International Conference on Information Security, Brighton, UK, May 1991.

Penzias, A., *"Voice Lock" Key to Future Security*, MIS Week, 11:11, Mar 12, 1990, p. 34.

Pfleeger, Charles P., Security in Computing, Prentice-Hall, 1989.

Porter, Joseph H. and Hamm, Robert J., *Statistics: Applications for the Behavioral Sciences*, Brooks/Cole Publishing Co., Monterey, CA, 1986.

Salamone, Salvatore, *Hard-to-Guess Passwords a Critical Key to Security*, Network World, 8:42, Oct 21, 1991, pp. 21-22.

Sawyer, Darren A., *The Characteristics of User-Generated Passwords*, Master's Thesis – Naval Postgraduate School, March 1990.

Smith, Gerald, *Selecting Secure Passwords*, Netware Advisor, 4:1, Jan 1991, p. 13.

Smith, S.L., *Authenticating Users by Word Association*, Computers and Security, 6:6, 1987, pp. 464-470.

Wood, Lamont, *Systems Security Using Token-Based Authentication*, Datamation, 37:14, Jul 15, 1991, pp. 69-70.

Zviran, Moshe and Haga, William J., *Cognitive Passwords: The Key to Easy Access Control*, Computers and Security, Vol. 9, 1990, pp. 723-736.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, VA 22304-6145	2
2.	Superintendent Attn: Library, Code 0142 Naval Postgraduate School Monterey, CA 93943	2
3.	Professor Moshe Zviran (Code AS/Zv) Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93943	1
4.	Professor William Haga (Code AS/Hg) Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93943	1
5.	Professor Tung Bui (Code AS/Bd) Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93943	1
6.	LT Kris Davis, USN Student Mailing Center Box 1412 Naval Postgraduate School Monterey, CA 93943	1
7.	LT Timothy Pence, USN COMSEVENTHFLT N2 Unit 25104 FPO AP 96601-6003	2